COMMUNICATION SCIENCES AND ENGINEERING

X. STATISTICAL COMMUNICATION THEORY*

Prof. Y. W. Lee
Prof. A. G. Bose
Prof. D. J. Sakrison
Prof. M. Schetzen
Prof. H. L. Van Trees, Jr.
V. R. Algazi
A. A. Al-Shalchi
R. Alter
D. S. Arnstein
M. E. Austin
R. F. Bauer
J. D. Bruce
A. M. Bush
J. K. Clemens
C. J. R. Deal
A. F. Gann
E. P. Gould
C. E. Gray
H. Hemami
W. F. Kelly
P. L. Konop
V. S. Mummert
R. B. Parente
T. L. Stewart
L. C. Thomas
W. S. Widnall

A. WORK COMPLETED

The following studies have been completed by the students indicated, and have been submitted as theses in partial fulfillment of the requirements for the degree of Master of Science, Department of Electrical Engineering, M.I.T., August 1962.

1. C. J. R. Deal, Optimum Laguerre Series Expansion of Functions.
2. T. L. Stewart, Measurement of Second-Order Correlation Functions by Orthogonal Expansions.
3. E. P. Gould, A Study of Error Criteria for Square-Law Devices.
4. A. A. Al-Shalchi, Prediction of Singular Time Functions.
5. R. D. Shelton, Estimation of Signal-to-Noise Ratio.

M. Schetzen

6. A Functional Study of Duffing's Equation

The present study has been completed by V. S. Mummert and the results have been presented to the Department of Aeronautics and Astronautics, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science, August 1962.

H. L. Van Trees, Jr.

B. A TWO-STATE MODULATION SYSTEM

A two-state modulation system for which the block diagram is shown in Fig. X-1 was described in Quarterly Progress Report No. 66 (pages 187-189). The present report presents a portion of the analysis of this modulation system to dc signals. The analysis presented is also valid for ac signals, provided that their highest frequency is low compared with the lowest switching frequency of the system.

The object of the present analysis is to determine the time average $\overline{y}$ of the output y(t) and the period T_2 of the modulator output as functions of the dc input signal E_s. The analysis is facilitated by reference to the sketch of the feedback signal e_{fb}

*This work was supported in part by the National Institutes of Health (Grant MH-04737-02).

shown in Fig. X-2.

We shall obtain expressions involving T_1, T_2, the system parameters, and E_s. From these expressions we can then determine $\bar{y}$.

First, let us find E_1 and then see how long it takes for e_{fb} to go from E_1 to $-E_s+\delta_2$. From Fig. X-2 we can write

$$E_1 = E_a + (-E_s+\delta_1-E_a)\, e^{-T_d/\tau}. \tag{1}$$

Next, we write an equation relating E_1 to $-E_s+\delta_2$ in the form

$$-E_s + \delta_2 = E_b + (E_1-E_b)\, e^{-(T_1-T_d)/\tau}. \tag{2}$$

Inserting Eq. 1 into Eq. 2 to eliminate E_1, we obtain

$$-E_s + \delta_2 - E_b = \left[E_a - E_b + (-E_s+\delta_1-E_a)\, e^{-T_d/\tau}\right] e^{-(T_1-T_d)/\tau}, \tag{3}$$

which relates T_1 to E_s and the system parameters. In a similar way two equations can be written which relate $-E_s+\delta_2$ to E_2 and E_2 to $-E_s+\delta_1$. When E_2 is eliminated from those equations, we obtain the result

$$-E_s + \delta_1 - E_a = \left[E_b - E_a + (-E_s+\delta_2-E_b)\, e^{-T_d/\tau}\right] e^{-(T_2-T_1-T_d)/\tau}. \tag{4}$$

The output signal $\bar{y}$ is related to T_1 and T_2 by the equation

$$\bar{y} = \frac{T_1}{T_2}(E_b-E_a) + E_a. \tag{5}$$

Equations 3, 4, and 5 can be solved simultaneously to yield the system transfer function and the period T_2 of the modulated signal. The results are

$$\bar{y} = \frac{\left\{\ln(-E_s+\delta_2-E_b) - \ln\left[E_a - E_b + (-E_s+\delta_1-E_a)\, e^{-T_d/\tau}\right] - \frac{T_d}{\tau}\right\}(E_b-E_a)}{\ln(-E_s+\delta_2-E_b) - \ln\left[E_a - E_b + (-E_s+\delta_1-E_a)\, e^{-T_d/\tau}\right] + \ln(E_a+E_s-\delta_1) - \ln\left[E_a - E_b + (E_b+E_s-\delta_2)\, e^{-T_d/\tau}\right] - \frac{2T_d}{\tau}} + E_a \tag{6}$$

and

$$T_2 = 2T_d + \tau \ln \frac{\left[E_a - E_b + (-E_s+\delta_1-E_a)\, e^{-T_d/\tau}\right]\left[E_a - E_b + (E_s-\delta_2+E_b)\, e^{-T_d/\tau}\right]}{(-E_s+\delta_2-E_b)(E_s-\delta_1+E_a)} \tag{7}$$

These general expressions are useful in determining the sensitivity of $\bar{y}$ and T_2 to dissymmetries in the hysteresis loop caused by drift in the circuitry that realizes the

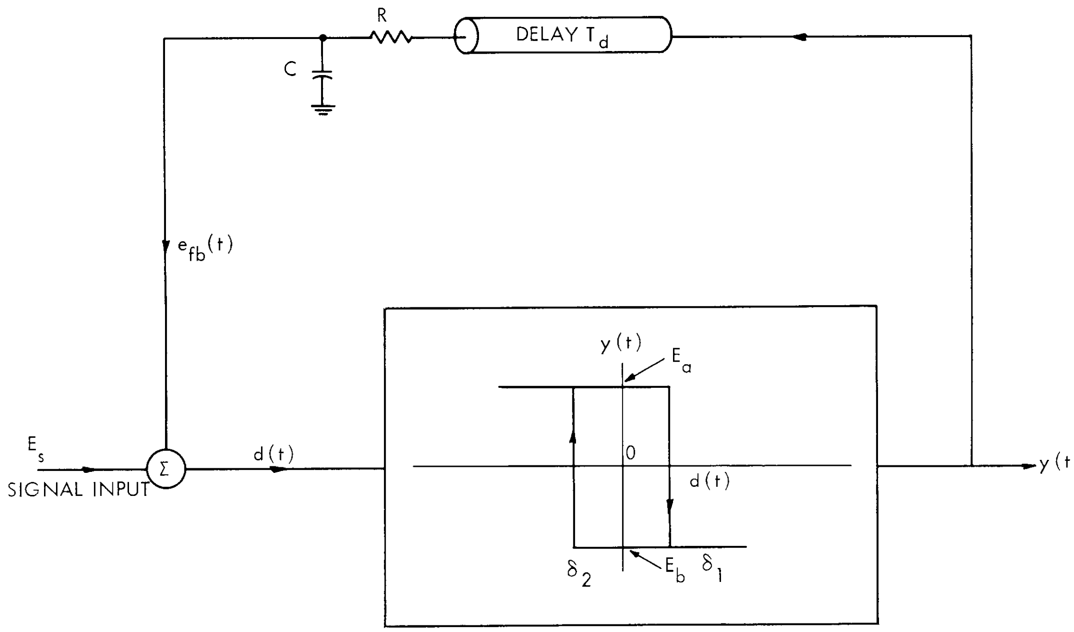

Fig. X-1. Diagram of modulation system indicating the system parameters.

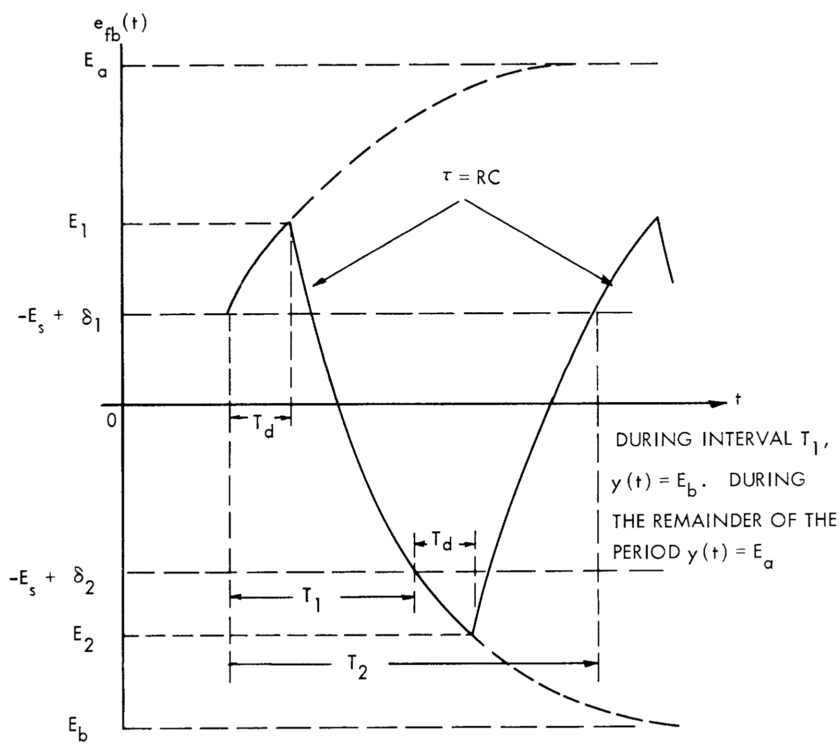

Fig. X-2. Feedback signal $e_{fb}(t)$.

loop. However, in order to obtain a better understanding of how $\overline{y}$ and T_2 depend upon the input signal E_s it is convenient and reasonable to assume a symmetrical hysteresis loop. Thus in Eqs. 6 and 7 we shall let $E_a = -E_b = E_o$ and $\delta_1 = -\delta_2 = \delta$. If we also normalize our expressions with respect to E_o, we obtain

$$\frac{\overline{y}}{E_o} = 1 - \cfrac{2}{1 + \cfrac{\ln\left[\dfrac{1 + E_s/E_o - \delta/E_o}{2 - (1-E_s/E_o-\delta/E_o)\, e^{-T_d/\tau}}\right] - \dfrac{T_d}{\tau}}{\ln\left[\dfrac{1 - E_s/E_o - \delta/E_o}{2 - (1+E_s/E_o-\delta/E_o)\, e^{-T_d/\tau}}\right] - \dfrac{T_d}{\tau}}} \tag{8}$$

and

$$T_2 = 2T_d + \tau \ln \frac{\left[2 - (1-\delta/E_o)\, e^{-T_d/\tau}\right]^2 - (E_s/E_o)^2\, e^{-2T_d/\tau}}{(1-\delta/E_o)^2 - (E_s/E_o)^2}. \tag{9}$$

The appearance of Eq. 8 does not at first glance indicate that $\overline{y}$ is a very linear function of E_s. However, a computer study of this relation for various values of the parameters δ/E_o and T_d/τ indicates that for practical values of the system parameters the system function is quite linear. For example, the ratio of output to input of the system is tabulated below for a range of input voltages and for the choice of parameters $\delta/E_o = 10^{-4}$ and $T_d/\tau = 0.281$.

E_s/E_o	$\overline{y}/E_s$	Percentage of Error
0.7	0.80225	1.532
0.35	0.78966	0.2732
0.175	0.78756	0.0631
0.0875	0.78708	0.0148
0.04375	0.78696	0.0029
0.021875	0.78693	0.0000
0.010937	0.78692	
0.54687×10^{-2}	0.78692	
0.27344×10^{-2}	0.78691	
0.13672×10^{-2}	0.78692	
0.68359×10^{-3}	0.78693	

The percentage of error column in this tabulation is the percentage by which $\overline{y}/E_s$ deviates from a straight line with a slope of 0.78693 at each value of E_s/E_o.

Since the primary objective of the present report is to derive expressions for the dc transfer function and the period of the modulation system, we shall not discuss at this time how we select the parameters or what effect this selection has upon the dynamic response. The choice of parameters above corresponds to a ratio of signal bandwidth to zero-signal modulation frequency of approximately 0.16. The linearity improves as this ratio is decreased.

Inspection of Eq. 9 reveals that the modulator period is a monotonically increasing function of E_s. For the choice of system parameters indicated above the modulator switching period increases from its zero-signal value by approximately 47 per cent for $E_s/E_o = 0.7$. Equation 9 shows that T_2 becomes infinite when $E_s/E_o = 1 - \delta/E_o$, a result that is easily understood, since for $E_s/E_o \geq 1 - \delta/E_o$ the feedback signal cannot attain a magnitude that is great enough to cause the system to switch states. The behavior of T_2 as a function of E_s suggests that we restrict the range of E_s in order to restrict the range of the switching frequency to frequencies well above the highest signal component.

A. G. Bose

C. DESIGN OF FILTERS FOR NON MEAN-SQUARE-ERROR PERFORMANCE CRITERIA BY MEANS OF A CONTINUOUS ADJUSTMENT PROCEDURE

In Quarterly Progress Report No. 66 (page 196), the author stated that he would give the proofs of Statements 1 and 2 concerning the convergence of the adjustment procedure in Quarterly Progress Report No. 67. The following report gives these proofs.

1. Proof of Statements 1 and 2

Note that in defining the adjustment process up to time t, our basic sample space is the space of functions $d(\tau)$, $f_1(\tau)$, ..., $f_k(\tau)$, $1 \leq \tau \leq t$. The quantity $\underline{x}_t$ is a random variable that depends not only on these functions but also on the deterministic quantity $\underline{x}(1)$. The quantity $\underline{x}_t$ is thus in reality a family of random variables indexed by the parameter $\underline{x}(1)$. When we state an equality or inequality, it is to be understood that it holds for all $\underline{x}(1) \in X$. Our starting point is the equation

$$E\left\{2(\underline{x}_t-\underline{\theta}, -a(t)/c(t)\underline{Y}_t) \mid \underline{x}_t\right\} = -a(t)\ 2(\underline{x}_t-\underline{\theta}, \underline{M}_{c(t)}(\underline{x}_t))$$

$$+ a(t)\cdot 2E\left\{\left(\underline{x}_t-\underline{\theta}, -\frac{1}{c(t)}\underline{Y}_t+M_{c(t)}(\underline{x}_t)\right) \middle| \underline{x}_t\right\}. \tag{1}$$

Here, $\underline{Y}_t$ is the vector random variable whose i^{th} component corresponds to the sample function $Y_i(t)$; $\underline{M}_{c(t)}(\underline{x}_t)$ denotes the expectation of $\underline{Y}_t(\underline{x}_t)$, in which $\underline{x}_t$ is treated as a vector and not as a random variable; that is, the quantity obtained by taking a long time average of $Y_t(x_t)$ with x and c held constant. Thus $\frac{1}{c(t)}\underline{M}_{c(t)}(\underline{x})$ is a symmetric difference approximation of $\operatorname{grad} M(\underline{z})|_{\underline{z}=\underline{x}}$. We consider separately the two terms on the right side of Eq. 1. We have shown[1] that with the random variables $f_{i,t}$ bounded, restriction (c) implies

$$P\left\{\left|\sum_{i=1}^{k} (x_i-\theta_i)\, f_{i,t}\right| \geq D\|\underline{x}-\underline{\theta}\|\right\} \geq \epsilon \tag{2}$$

for $D > 0$, $\epsilon > 0$, $\underline{x} \in X$, $\underline{x} \neq \underline{\theta}$. It has also been shown[2,3] that restrictions (a), (b), (d) and inequality (2) imply that

$$(\underline{x}-\underline{\theta}, \operatorname{grad} M(\underline{z})|_{\underline{z}=\underline{x}}) \geq k_o\|\underline{x}-\underline{\theta}\|^2 \qquad k_o > 0\,. \tag{3}$$

It has also been shown[4] that

$$\underline{M}_c(\underline{x}) = 2 \operatorname{grad} M(\underline{z})|_{z=\underline{x}} + c^2/6\,\underline{z} \tag{4}$$

in which $\underline{z}$ is a vector whose i^{th} component is

$$\left.\frac{\partial^3 M(\underline{z})}{\partial x_i^3}\right|_{z=\underline{x}+c\tau_i\underline{e}_i} - \left.\frac{\partial^3 M(\underline{z})}{\partial x_i^3}\right|_{\underline{z}=\underline{x}-c\tau_i'\underline{e}_i} \qquad 0 \leq \tau_i \leq 1 \qquad 0 \leq \tau_i' \leq 1\,.$$

Combining Eq. 4 and inequality (3), we obtain

$$-(\underline{x}-\underline{\theta}, \underline{M}_c(\underline{x})) \leq -2k_o\|\underline{x}-\underline{\theta}\|^2 + (k)^{1/2}\frac{c^2}{3} Q\|\underline{x}-\underline{\theta}\| \tag{5}$$

in which

$$Q = \sup_{\underline{x}\in X\ i=1,2,\ldots,k} \left|\left.\frac{\partial^3 M(z)}{\partial x_i^3}\right|_{z=\underline{x}}\right|\,.$$

This quantity Q can be shown to be bounded by restrictions (a), (b), and (d). Substituting inequality (5) in Eq. 1 and taking expected values, we obtain

$$E\left\{2\left(\underline{x}_t-\underline{\theta}, -\frac{a(t)}{c(t)}\underline{Y}_t\right)\right\} \leq -4k_o a(t)\, E\{\|\underline{x}_t-\underline{\theta}\|^2\} + \frac{2}{3}a(t)\, c^2(t)\, k^{1/2} Q E\{\|\underline{x}_t-\underline{\theta}\|\}$$
$$+ 2a(t)\left|E\left\{\left(\underline{x}_t-\underline{\theta}, -\frac{1}{c(t)}\underline{Y}_t+\underline{M}_{c(t)}(\underline{x}_t)\right)\right\}\right|\,. \tag{6}$$

It will be shown that restrictions (a), (b), (d), and (e) imply that

$$\left|E\left\{\left(\underline{x}_t-\underline{\theta}, -\frac{1}{c(t)}\underline{Y}_t+\underline{M}_{c(t)}(\underline{x}_t)\right)\right\}\right| \leq k_3 a(t/2)\,. \tag{7}$$

We now have satisfactory bounds on the right-hand side of inequality (6). Consideration of the left-hand side of this inequality reveals that it is equal to

$$E\left\{\frac{d}{dt}\|\underline{x}_t-\underline{\theta}\|^2\right\}.$$

Restrictions (ii), (a), (b), and (d) guarantee that we can interchange the order of differentiation and expectation[5] in this quantity. Thus, letting b(t) denote $E\{\|\underline{x}_t-\underline{\theta}\|^2\}$ and substituting inequality (7) in inequality (6), we obtain

$$\frac{d}{dt}\,b(t) \leq -4k_o a(t)\,b(t) + \frac{2}{3}\,a(t)\,c^2(t)\,k^{1/2}QE\{\|\underline{x}_t-\underline{\theta}\|\} + 2K_3 a(t)\,a(t/2). \tag{8}$$

It is easily shown that

$$E\{\|\underline{x}_t-\underline{\theta}\|\} \leq \epsilon_t + \frac{1}{\epsilon_t}\,E\{\|\underline{x}_t-\underline{\theta}\|^2\} \tag{9}$$

for any $\epsilon_t > 0$. Setting

$$\epsilon_t = \frac{2}{3}\,\frac{k^{1/2}Qc^2(t)}{\epsilon K_o} \qquad 0 < \epsilon < 4$$

and substituting inequality (9) in inequality (8), we obtain

$$\frac{d}{dt}\,b(t) + p(t)\,b(t) \leq q(t) \tag{10}$$

in which

$$q(t) = \frac{4}{9}\,\frac{k^2Q}{\epsilon K_o}\,a(t)\,c^4(t) + 2k_3 a(t)\,a\left(\frac{t}{2}\right) > 0$$

$$p(t) = (4-\epsilon)\,K_o a(t) > 0.$$

Integrating both sides of inequality (10) from 1 to t and rearranging terms, we have

$$b(t) + \int_1^t p(\tau)\,b(\tau)\,d\tau \leq \int_1^t q(\tau) + b(1). \tag{11}$$

Now consider the integral equation

$$b_o(t) + \int_1^t p(\tau)\,b(\tau)\,d\tau = \int_1^t q(\tau)\,d\tau + b(1) \tag{12}$$

with the solution

$$b_o(t) = \left[\exp -\int_1^t p(\tau)\,d\tau\right]\left[\int_1^t q(\tau)\,\exp\int_1^\tau p(\zeta)\,d\zeta\,d\tau + b(1)\right]. \tag{13}$$

It can be shown in a straightforward manner that, for b(t) satisfying inequality (11) and $b_o(t)$ satisfying Eq. 12,

$$b(t) \leq b_o(t), \tag{14}$$

and thus we shall concentrate on bounding $b_o(t)$. Rewriting Eq. 13, we have

$$b_o(t) = b(1) \exp\left[-\int_1^t p(\tau)\, d\tau\right] + \int_1^\infty f_{[1,t]}(\tau)\, d\tau \tag{15}$$

in which

$$f_{[1,t]}(\tau) = \begin{cases} \left[\exp-\int_\tau^t p(\xi)\, d\xi\right] q(\tau) & 1 \leq \tau \leq t \\ 0 & \text{elsewhere} \end{cases} \tag{16}$$

By restriction (1) it follows that

$$0 \leq f_{[1,t]}(\tau) \leq q(\tau) \qquad \text{for all } t \text{ and } \tau \text{ greater than one}$$

and

$$\int_1^t q(\tau)\, d\tau \qquad \text{exists for all } t \geq 1.$$

Moreover, from Eq. 16

$$\lim_{t\to\infty} f_{[1,t]}(\tau) = 0 \qquad \text{for all } \tau. \tag{17}$$

Therefore, by the general convergence theorem of Lebesque[6]

$$\lim_{t\to\infty} b_o(t) = \lim_{t\to\infty} b(1) \exp\left[-\int_1^t p(\tau)\, d\tau\right] + \lim_{t\to\infty} \int_1^\infty f_{[1,t]}(\tau)\, d\tau$$

$$= 0 + \int_1^\infty \lim_{t\to\infty} f_{[1,t]}(\tau)\, d\tau = 0. \tag{18}$$

The first term on the right-hand side of Eq. 18 goes to zero because b(1) is bounded and the integral of $p(\tau)$ is divergent. This equation, together with inequality (14), completes the proof of Statement 1.

To complete the proof of the positive side of Statement 2, we merely note that for a(t) and c(t), as specified in Statement 2,

$$q(t) \leq k_4 \frac{1}{t^2}, \quad p(t) = (4-\epsilon) k_o \frac{a}{t} > 0, \quad t > 1.$$

Substituting these in Eq. 13 and simplifying, we obtain

$$b_o(t) \leqslant b(1) \frac{1}{t^{(4-\epsilon)K_o a}} + K_4[(4-\epsilon)K_o a-1] \frac{1}{t} - \frac{1}{t^{(4-\epsilon)K_o a}} \qquad 0 < \epsilon < 4. \tag{19}$$

Thus, noting that b(1) is bounded, we have, for $a > \frac{1}{4k_o}$

$$b_o(t) \leqslant K_5 \frac{1}{t}. \tag{20}$$

The negative side of Statement 2 is established by referring to inequality (8) and recalling that the last term $2k_3 a(t)\, a(t/2)$ resulted from finding an upper bound for the quantity

$$C = E\left\{\left[\underline{x}_t - \underline{\theta}, -\frac{1}{c(t)}\underline{Y}_t + \underline{M}_{c(t)}(\underline{x}_t)\right]\right\}. \tag{21}$$

By picking a suitable example ($W(e) = e^2$ and d_t and f_t possessing simple suitable statistical properties consistent with restrictions (a)-(e)) it is possible to obtain a lower bound for C which is of the form

$$C \geqslant k_6 a\left(\frac{t}{2}-1\right) \qquad k_6 > 0.$$

This will be the dominant term in inequality (8), since for $W(e) = e^2$ the quantity Q is zero. Under these conditions, it is possible to use a line of reasoning analogous to that above to show that

$$b(t) \geqslant k_7 \frac{1}{t} \qquad k_7 > 0.$$

This establishes the positive side of Statement 2 and thus completes the proof.

2. Proof of a Certain Inequality

We wish to establish inequality (7) by suitably bounding

$$|C| = \left|E\left\{\left(\underline{x}_t - \underline{\theta}, -\frac{1}{c(t)}\underline{Y}_t + \underline{M}_{c(t)}(\underline{x}_t)\right)\right\}\right|. \tag{22}$$

Here, we treat only the one-dimensional case: $\underline{x} = x$, a single parameter. The method and most of the details are the same for the multidimensional case; the only additional problem is that of notation.

We have

$$\frac{1}{c(t)} Y_t = \frac{1}{c(t)} \{W[d_t-(x_t+c(t)f_t]-W[d_t-(x_t-c(t))f_t]\}.$$

Recognizing that W is a finite polynomial of order N, expanding and collecting terms, we obtain

$$\frac{1}{c(t)} Y_t = \frac{1}{c(t)} \left\{\sum_{j=0}^{N} \sum_{m=0}^{j} A_{jm}(d_t)^{N-j}(f_t)^m\left[(x_t+c(t))^m-(x_t-c(t))^m\right]\right\}. \tag{23}$$

For m = 0 the argument of the double sum is zero, and for $m \geq 1$

$$(x_t+c(t))^m - (x_t-c(t))^m = 2 \sum_{\substack{q=1 \\ q \text{ odd}}}^{m} \binom{m}{q} (x_t)^{m-q} (c(t))^q . \tag{24}$$

Substituting Eq. 24 in Eq. 23, we obtain

$$\frac{1}{c(t)} Y_t = \sum_{j=0}^{N} \sum_{m=1}^{j} \sum_{\substack{q=1 \\ q \text{ odd}}}^{m} D_{jmq}(c(t))^{q-1} (x_t)^{m-q} (d_t)^{N-j} (f_t)^m . \tag{25}$$

Moreover, by our definition of $M_{c(t)}(x_t)$

$$M_{c(t)}(x_t) = \sum_{j=0}^{N} \sum_{m=1}^{j} \sum_{\substack{q=1 \\ q \text{ odd}}}^{m} D_{jmq}(x_t)^{m-q} (c(t))^{q-1} \overline{(d_t)^{N-j} (f_t)^m} . \tag{26}$$

Substituting Eqs. 25 and 26 in Eq. 27, we obtain

$$|C| = \left| \sum_{j=0}^{N} \sum_{m=1}^{j} \sum_{\substack{q=1 \\ q \text{ odd}}}^{m} - D_{jmq}(c(t))^{q-1} \overline{(x_t-\theta)(x_t)^{m-q}\left[(d_t)^{N-j}(f_t)^m - \overline{(d_t)^{N-j}(f_t)^m}\right]} \right. . \tag{27}$$

in which the bar indicates expectation. This expression consists of a finite number of terms, the number being dependent only on N, the degree of W. We shall consider a typical term of this expression and bound it in magnitude. We pick a term in which c(t) does not appear, since these terms approach zero at a slower rate than those in which c(t) is raised to a positive power. We also pick a term that does not contain θ; a term containing θ can be handled in the same fashion.

We thus wish to bound

$$|C_r| = \left|\overline{(x_t)^r(\zeta_t-\overline{\zeta}_t)}\right| \tag{28}$$

in which

$$\zeta_t = (d_t)^s(f_t)^p . \tag{29}$$

We may rewrite Eq. 28 as

$$C_r = \overline{(x_t)^r(\zeta_t-\overline{\zeta}_t)} = (x(1))^r\overline{(\zeta_t-\overline{\zeta}_t)} + \int_1^t \overline{(\zeta_t-\overline{\zeta}_t)\left[\frac{d}{d\tau}(x_\tau)^r\right]}\, d\tau$$

$$= \int_1^t \overline{(\zeta_t-\overline{\zeta}_t)\left[\frac{d}{d\tau}(x_\tau)^r\right]}\, d\tau \tag{30}$$

in which we are justified in interchanging the order of integration and expectation by the theorem of Kolmogoroff.[5] Now

$$\frac{d}{d\tau}(x_\tau)^r = r(x_\tau)^{r-1}\frac{d}{d\tau}(x_\tau) = r(x_\tau)^{r-1}\left[-\frac{a(\tau)}{c(\tau)}\right]Y_\tau . \tag{31}$$

Combining Eqs. 28-31, we obtain

$$C_r = -r\int_1^t a(\tau)\,\overline{(x_\tau)^{r-1}\left(\frac{1}{c(\tau)}Y_\tau\right)(\zeta_t-\overline{\zeta}_t)}\,d\tau . \tag{32}$$

Refer to Eq. 25; both x_τ and $\frac{1}{c(\tau)}Y_\tau$ are bounded random variables that depend on f(s) and d(s) only for $s \leqslant \tau$. Thus by restriction (e)

$$\left|\overline{(x_\tau)^{r-1}\left(\frac{1}{c(\tau)}Y_\tau\right)(\zeta_t-\overline{\zeta}_t)}\right| \leqslant \begin{cases} K_5 & t-1 \leqslant \tau \leqslant t \\ \dfrac{K_5}{(t-\tau)^2} & \tau \leqslant t-1 \end{cases} \tag{33}$$

$$K_5 < \infty \qquad \text{for all } x(1) \in X .$$

Considering restriction (i), we shall assume that all functions a(t) of interest will be monotonic. Thus, substituting inequality (33) in Eq. 32, we obtain

$$\begin{aligned} |C_r| &\leqslant rK_5\left[\int_1^{t-1} a(\tau)\frac{1}{(t-\tau)^2}\,d\tau + \int_{t-1}^{t} a(\tau)\,d\tau\right] \\ &\leqslant rK_5\left\{a\left(\frac{t}{2}\right)\left[1+\int_{t/2}^{t-1}\frac{1}{(t-\tau)^2}\,d\tau\right] + a(1)\int_1^{t/2}\frac{1}{(t-\tau)^2}\,d\tau\right\} \\ &\leqslant rK_5\left[2a\left(\frac{t}{2}\right)+a(1)\frac{2}{t}\right] \qquad \text{all } x(1) \in X . \end{aligned} \tag{34}$$

But, by assuming a(t) to be monotonic, restriction (i) implies that $a(t/2) \geqslant K_6\,1/t$, $K_6 > 0$, and

$$|C_r| \leqslant K_7 r a\left(\frac{t}{2}\right), \quad K_7 < \infty, \qquad \text{all } x(1) \in X . \tag{35}$$

We have thus split C into a finite number of terms and shown that each of these can be bounded in magnitude as in inequality (35). Thus

$$|C| \leqslant K_3 a\left(\frac{t}{2}\right), \quad K_3 < \infty, \qquad \text{for all } x(1) \in X$$

as desired.

D. J. Sakrison

References

1. D. J. Sakrison, Application of Stochastic Approximation Methods to System Optimization, Technical Report 391, Research Laboratory of Electronics, M.I.T., July 10, 1962, pp. 70-72.

2. D. Sakrison, Application of stochastic approximation methods to optimum filter design, IRE International Convention Record, Part 4, pp. 127-136, 1961.

3. D. J. Sakrison, Application of Stochastic Approximation Methods to System Optimization, op. cit., pp. 19-26.

4. Ibid., pp. 59-63.

5. A. M. Kolmogoroff, Foundations of the Theory of Probability (Chelsea Publishing Company, New York, 1956), p. 44.

6. E. C. Titchmarsh, The Theory of Functions (Oxford University Press, London, 2d edition, 1939), p. 345.

D. PREDICTION OF SINGULAR TIME FUNCTIONS

In pure prediction the desired output of the predictor is the input advanced by the prediction time, α seconds. That is,

$$f_d(t) = f_i(t+\alpha) . \tag{1}$$

The optimum realizable linear system for a minimum mean-square-error criterion is given by the Wiener-Hopf equation[1]

$$\phi_{ii}(\tau+\alpha) = \int_0^\infty h_{opt}(\sigma)\, \phi_{ii}(\tau-\sigma)\, d\sigma \qquad \text{for } \tau \geq 0 , \tag{2}$$

in which $\phi_{ii}(\tau)$ is the autocorrelation function of the input, $f_i(t)$. If the power density spectrum, $\Phi_{ii}(\omega)$, of the input is factorable, it can be written as

$$\Phi_{ii}(\omega) = \Phi_{ii}^{+}(\omega)\, \Phi_{ii}^{-}(\omega) , \tag{3}$$

in which $\Phi_{ii}^{+}(\omega)$ is the complex conjugate of $\Phi_{ii}^{-}(\omega)$; also, all of the poles and zeros of $\Phi_{ii}^{+}(\omega)$ are in the left-half of the complex S-plane in which $S = \sigma + j\omega$. In terms of $\Phi_{ii}^{+}(\omega)$, the solution of the Wiener-Hopf equation is

$$H_{opt}(\omega) = \frac{1}{2\Pi\Phi_{ii}^{+}(\omega)} \int_0^\infty \psi(\tau+\alpha)\, e^{-j\omega\tau}\, d\tau , \tag{4}$$

in which

$$\psi(\tau) = \int_{-\infty}^{\infty} \Phi_{mm}^{+}(\omega)\, e^{j\omega\tau}\, d\omega .$$

The minimum mean-square error in pure prediction is then

$$\overline{\mathscr{E}^2_{h_{opt}}(t)} = \frac{1}{2\Pi} \int_0^{\alpha} \psi^2(\tau)\, d\tau . \tag{5}$$

It is seen that the mean-square error is a monotonically increasing function of the prediction time, α.

The basis of this method for determining the optimum linear predictor is the factorization of $\Phi_{ii}(\omega)$. There are, however, functions for which this factorization is not possible. Wiener has shown that factorization is not possible for those power density spectra for which

$$\int_{-\infty}^{\infty} \frac{|\ln \Phi_{ii}(\omega)|}{1+\omega^2}\, d\omega = \infty . \tag{6}$$

This integral is called the Paley-Wiener integral.[2] Furthermore, Wiener has shown that if the power density spectrum, $\Phi_{ii}(\omega)$, is not factorable, the future of the time function, $f_i(t)$, can be determined completely by a linear operation on its past. That is, for any prediction time the mean-square error of prediction with a realizable linear system can be made arbitrarily small. In communication the future of all practical messages is not completely determined by its own past, since it would then not be possible to introduce new information in any period of the message. Thus time functions for which the Paley-Wiener integral diverges are called singular time functions.

In this report we shall present some results that we have obtained concerning the properties of these singular time functions which enable them to be predicted with an arbitrarily small error by means of a realizable linear system.

1. Classes of Singular Functions

We find it convenient to classify singular time functions in two classes in accordance with the manner by which they cause the Paley-Wiener integral to diverge. The first class, which we denote Class A, is made up of exponentially band-restricted time functions. These are time functions whose power density spectra are, for large ω, exponentially bounded. That is, for $|\omega| \geq W$

$$\Phi_{ii}(\omega) \leq B\, e^{-|A\omega|} . \tag{7}$$

The second class, which we denote Class B, is made up of band-eliminated time functions. These are time functions that contain no power over at least a band of frequencies. That is, their power density spectra are zero in some band of frequencies. It is, of course, possible for a time function to be a member of both classes simultaneously. An example of such a function is the bandlimited time function that contains no power above a certain frequency. We shall first discuss the prediction of functions of Class A. The results obtained for this class will then be extended to the prediction of functions of Class B.

2. The Prediction of Exponentially Band-Restricted Functions

Since power density spectra for functions of Class A cause the Paley-Wiener integral (Eq. 6) to diverge, they are not factorable and Eqs. 4 and 5 are not applicable. Insight into a desired transfer function of the predictor can be obtained from a general expression for the mean-square error in prediction. The desired expression is obtained by noting that the desired output of the predictor as given by Eq. 1 can be thought of as the response of an unrealizable linear system whose impulse response is an impulse at $t = -\alpha$. The transfer function of such a system is $e^{j\omega\alpha}$. If a linear system with the transfer function $H(\omega)$ is used to predict the signal $f_i(t)$, then the error can be considered the output of the linear system depicted in Fig. X-3. The transfer function of this system is

$$G(\omega) = e^{j\omega\alpha} - H(\omega). \tag{8}$$

Then, since

$$\overline{\mathscr{E}^2(t)} = \phi_{ee}(0) = \int_{-\infty}^{\infty} \Phi_{ee}(\omega)\, d\omega, \tag{9}$$

we have

$$\begin{aligned} \overline{\mathscr{E}^2(t)} &= \int_{-\infty}^{\infty} |G(\omega)|^2\, \Phi_{ii}(\omega)\, d\omega \\ &= \int_{-\infty}^{\infty} |e^{j\omega\alpha} - H(\omega)|^2\, \Phi_{ii}(\omega)\, d\omega. \end{aligned} \tag{10}$$

We note from Eq. 10 that for the mean-square error of prediction to be reducible to an arbitrarily small value, it is not necessary to make $G(\omega) = 0$. Rather, it is just

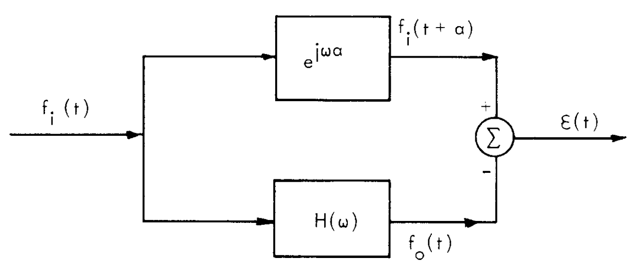

Fig. X-3. Pertaining to the calculation of the mean-square error of prediction.

necessary to make the area under the product $|G(\omega)|^2\ \Phi_{ii}(\omega)$ arbitrarily small. Thus, for example, for band-eliminated time functions, the value of $G(\omega)$ in the bands in which $\Phi_{ii}(\omega) = 0$ does not affect the mean-square error of prediction. This fact is the basis for our ability to reduce the mean-square error of prediction to an arbitrarily small value with a realizable linear system, $H(\omega)$. We also note that the optimum predictor for such time functions is not unique.

To determine a predictor for time functions of Class A, we note that their power density spectra are exponentially restricted for large ω. Thus, it is possible for $|G(\omega)|$ to increase with increasing ω and the mean-square error of prediction to be small; as a result, a possible form of a realizable linear predictor for such time functions is a series of differentiators. Thus we choose $H(\omega)$ to be the first N terms of the MacLaurin series of expansion of $e^{j\omega a}$. That is,

$$H(\omega) = \sum_{n=0}^{N-1} \frac{a^n}{n!} (j\omega)^n . \tag{11}$$

Then, from Eq. 8

$$G(\omega) = e^{j\omega a} - \sum_{n=0}^{N-1} \frac{a^n}{n!} (j\omega)^n . \tag{12}$$

By the use of the remainder theorem for the Taylor series, we than have

$$|G(\omega)| = \frac{|(a\omega)^N|}{N!} . \tag{13}$$

The mean-square error of prediction is obtained by substituting Eq. 13 in Eq. 10. It is

$$\overline{\mathscr{E}^2(t)} = \int_{-\infty}^{\infty} \frac{(a\omega)^{2N}}{(N!)^2} \Phi_{ii}(\omega)\, d\omega . \tag{14}$$

We shall consider the case for which

$$\Phi_{ii}(\omega) \leq e^{-|\omega|^k} \qquad \text{for } |\omega| \geq W, \tag{15}$$

first, for $k > 1$ and, second, for $k = 1$. For $k > 1$, an upper bound of the value of the mean-square error can then be obtained by use of inequality (15). Thus, from Eq. 14

$$\overline{\mathscr{E}^2(t)} = 2\int_0^W \frac{(a\omega)^{2N}}{(N!)^2} \Phi_{ii}(\omega)\, d\omega + 2\int_W^{\infty} \frac{(a\omega)^{2N}}{(N!)^2} \Phi_{ii}(\omega)\, d\omega$$

$$\leq \frac{(aW)^{2N}}{(N!)^2} 2\int_0^W \Phi_{ii}(\omega)\, d\omega + \frac{2a^{2N}}{(N!)^2} \int_W^{\infty} \omega^{2N} e^{-\omega^k}\, d\omega . \tag{16}$$

The first term of Eq. 16 can be made arbitrarily small by choosing N sufficiently large because

$$I_1 \leqslant \frac{(aW)^{2N}}{(N!)^2} \int_{-\infty}^{\infty} \Phi_{ii}(\omega)\, d\omega \leqslant \left[\frac{(aW)^N}{N!}\right]^2 \phi_{ii}(0)$$

and

$$\lim_{N\to\infty} \frac{(aW)^N}{N!} = 0 .$$

The second term of Eq. 16 can also be made arbitrarily small by choosing N sufficiently large. This fact can be shown as follows:

$$I_2 = \frac{2a^{2N}}{(N!)^2} \int_W^{\infty} \omega^{2N} e^{-\omega^k}\, d\omega \leqslant \frac{2a^{2N}}{(N!)^2} \int_0^{\infty} \omega^{2N} e^{-\omega^k}\, d\omega .$$

Now let $x = \omega^k$. Then

$$I_2 \leqslant \frac{2a^{2N}}{k(N!)^2} \int_0^{\infty} x^{\left(\frac{2N+1-k}{k}\right)} e^{-x}\, dx \leqslant \frac{2a^{2N}}{k(N!)^2} \Gamma\left(\frac{2N+1}{k}\right)$$

$$= \frac{2a^{2N}}{k(N!)^2} \left(\frac{2N+1-k}{k}\right)! < \frac{2a^{2N}}{k(N!)^2} \left(\frac{2N}{k}\right)! \tag{17}$$

By use of the Stirling approximation, we obtain

$$I_2 < 2 \sqrt{\frac{1}{\Pi k N}} \left(\frac{2a}{kN^{k-1}}\right)^{2N} \exp\left[2N\left(\frac{k-1}{k}\right)\right] \tag{18}$$

for sufficiently large N.

Thus, for $k > 1$, we can make I_2 arbitrarily small by choosing N sufficiently large because

$$\lim_{x\to\infty} \frac{e^{ax}}{x^{bx}} = 0 \qquad \text{for } a, b > 0 .$$

Thus the mean-square error of prediction can be made arbitrarily small by using a suitably large number of differentiators. That is, we have shown that

$$\lim_{N\to\infty} \overline{\left[f_i(t+a) - \sum_{n=0}^{N} \frac{a^n}{n!} \frac{d^n}{dt^n} f_i(t)\right]^2} = 0 \tag{19}$$

for time functions of Class A whose power density spectra are restricted by Eq. 15 for $k > 1$. We note the correspondence of this class of functions to analytic functions.

For our demonstration we required that $k > 1$. However, functions for which $k = 1$

are also members of Class A. It is not surprising that we have some difficulty with these functions because, in a sense, they form boundary functions between singular and nonsingular time functions. This result follows because time functions whose power density spectra for large ω decay to zero slower than $e^{-|A\omega|}$ are nonsingular since, for such spectra, the Paley-Wiener integral converges. To predict time functions for which

$$\Phi_{ii}(\omega) \leqslant B\, e^{-|A\omega|} \tag{20}$$

we need to be more subtle in our use of differentiators. The technique is to form a complete orthonormal set of functions by using differentiators. Using the Gram-Schmidt procedure, we form an orthonormal set of functions, $\{P_n(\omega) \exp[-1/2|A\omega|]\}$, from the set $\{F_n(\omega) \exp[-|A\omega|/2]\}$ in which

$$F_n(\omega) \exp\left[-\tfrac{1}{2}|A\omega|\right] = (j\omega)^n \exp\left[-\tfrac{1}{2}|A\omega|\right] . \tag{21}$$

The functions, $\{P_n(\omega) \exp[-1/2|A\omega|]\}$, thus formed are similar to the Laguerre functions. We now form the expansion

$$e^{j\omega a} \exp\left[-\tfrac{1}{2}|A\omega|\right] = \sum_{n=0}^{N} C_n P_n(\omega) \exp\left[-\tfrac{1}{2}|A\omega|\right] \tag{22}$$

in which

$$C_n = \int_{-\infty}^{\infty} e^{j\omega a}\, P_n(\omega)\, e^{-|A\omega|}\, d\omega , \tag{23}$$

since

$$\int_{-\infty}^{\infty} P_n(\omega) P_m(\omega)\, e^{-|A\omega|}\, d\omega = \begin{cases} 1 & \text{for } m = n \\ 0 & \text{for } m \neq n . \end{cases} \tag{24}$$

The integral-square error obtained by using N terms of this set is

$$\begin{aligned} \mathscr{E}_N &= \int_{-\infty}^{\infty} \left| e^{j\omega a} \exp\left[-\tfrac{1}{2}|A\omega|\right] - \sum_{n=0}^{N} C_n P_n(\omega) \exp\left[-\tfrac{1}{2}|A\omega|\right]\right|^2 d\omega \\ &= \int_{-\infty}^{\infty} \left| e^{j\omega a} - \sum_{n=0}^{N} C_n P_n(\omega)\right|^2 e^{-|A\omega|}\, d\omega . \end{aligned} \tag{25}$$

This error tends to zero as N increases, since the set is complete and

$$\int_{-\infty}^{\infty} \left| e^{j\omega a} \exp\left[-\tfrac{1}{2}|A\omega|\right]\right|^2 d\omega < \infty . \tag{26}$$

By substituting Eq. 20 in Eq. 10, we note that expression (25) for the integral-square error is an upper bound to the mean-square error of prediction. Thus, since $\lim_{N\to\infty} \mathscr{E}_N = 0$, the mean-square error of prediction can be made arbitrarily small by using a suitably large number of members of the set $\{P_n(\omega)\}$ as the predictor. We note that the functions $P_n(\omega)$ are a linear combination of differentiators. Thus we are still predicting the time function, $f_i(t)$, by means of differentiators; however, the coefficients – which correspond to amplifier gains – are different from those given in Eq. 19. We have thus shown that, for any prediction time, all functions of Class A can be predicted with arbitrarily small mean-square error by using only differentiators. This result implies that for any member of Class A the entire function – its complete past, as well as its complete future – is determined by the value of the function and all of its derivatives at one instant of time. Thus, for example, a bandlimited time function that has no power above W cycles per second could be specified in terms of its Taylor series expansion at some instant of time instead of its sample values spaced every 1/2W seconds apart.

3. Prediction of Band-Eliminated Time Functions

We shall study the prediction of this class by approximating spectra in this class by sequences of nonsingular spectra. The prediction of the nonsingular spectrum is known and its results are given by Eqs. 2-5. As the approximation becomes very close, we shall show that the mean-square error of prediction becomes very small. By virtue of Eq. 10, this result means that the optimum predictor approaches the perfect predictor arbitrarily closely over the nonzero bands of the spectrum.

Consider the band-eliminated spectrum, $\Phi_{aa}(\omega)$, shown in Fig. X-4a. It can be approximated by

$$\Phi_n(\omega) = \frac{\omega^{2n}}{1+\omega^{2n}}. \tag{27}$$

This $\Phi_n(\omega)$ is a nonsingular spectrum in the sense that its Paley-Wiener integral converges. With a little algebra one can see that, in the complex S-plane, the poles become smoothly distributed over the circle as n becomes large.

We shall find the optimum predictor for $\Phi_n(\omega)$ by the method of spectrum factorization. Let us express the spectrum as a product in terms of its poles, p_i, and its zeros, z_i. Then, factorizing it, we obtain

$$\Phi_n^+(S) = \prod_i \frac{(S-z_i)}{(S-p_i)} \tag{28}$$

where the zeros, z_i, and the poles, p_i, are simple and lie in the left-half plane. Next, we expand the product into partial fractions

$$\Phi_n^+(S) = \sum_k \frac{C_k}{(S-p_k)} \tag{29}$$

$$C_k = \prod_i (p_k - z_i) \Big/ \prod_{i,\, i\neq k} (p_k - p_i). \tag{30}$$

Using Eq. 4, we find the optimum predictor for prediction time a

$$H_{opt_n}(S) = \sum_k \frac{C_k e^{ap_k}/(S-p_k)}{\Phi_n^+(S)}$$

$$= \sum_k e^{ap_k} \prod_{i,\, i\neq k} \frac{S - p_i}{p_k - p_i} \prod_i \frac{p_k - z_i}{S - z_i}. \tag{31}$$

At the poles of the spectrum the optimum predictor is equal to the perfect predictor; however, at the zeros of the spectrum the difference, $G(\omega)$, between the optimum predictor and $e^{ja\omega}$ is large. Using Eq. 31 we find that the optimum predictor for $\Phi_n(\omega)$ is

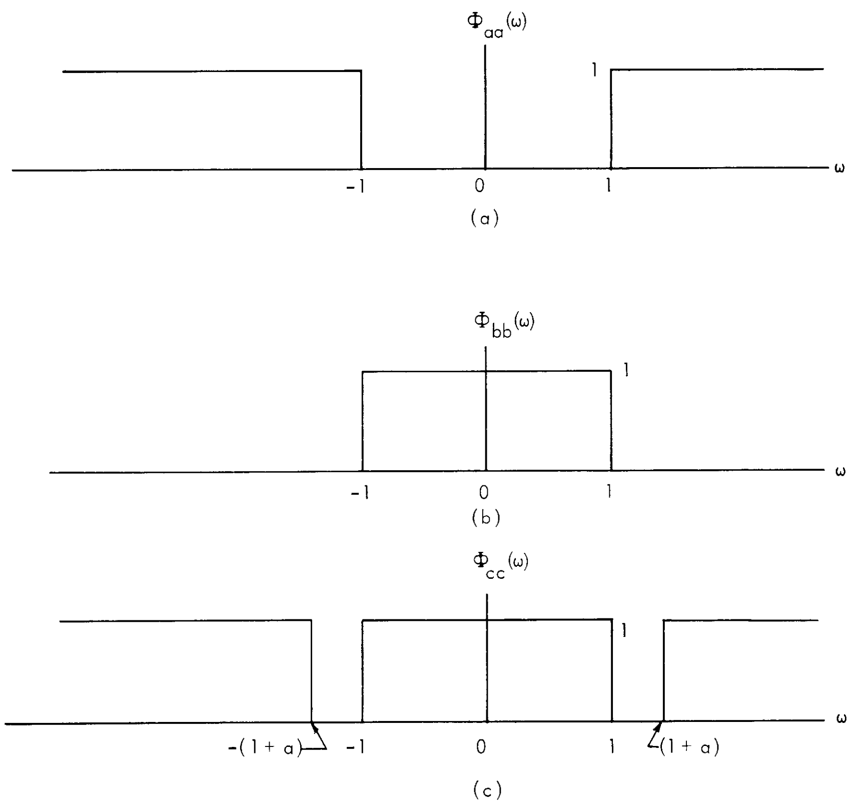

Fig. X-4. Some band-eliminated spectra.

$$H_{opt_n}(S) = \sum_{k=1}^{n} \left(\frac{S_k}{S}\right)^n e^{aS_k} \prod_{i=1,\, i\neq k}^{n} \frac{S - S_i}{S_i - S_k} \tag{32}$$

$$S_i = \pm j \exp\left[j\left(\frac{\Pi}{2n} + \frac{k\Pi}{n}\right)\right] \qquad k = 0, 1, \ldots, n, \tag{33}$$

and the S_i lie in the left-hand plane only. As n becomes very large, $\Phi_n(\omega)$ becomes very close to $\Phi_{aa}(\omega)$ and it can be shown that the optimum predictor converges to a perfect predictor.[3]

Note that the predictor is a series of integrators. These are large at small frequencies, at which our spectrum $\Phi_n(\omega)$ is small. Thus not much error is introduced. This result is similar to that for functions of Class A, which, as we have found from Eqs. 19 and 25, are predictable by differentiators because the spectrum is small at high frequencies, at which differentiators are large.

Next, we shall consider the behaviour of the mean-square error in predicting $\Phi_n(\omega)$ for very large n. The mean-square error of prediction $\overline{\mathscr{E}_n^2(t)}$ given by Eq. 5 is

$$\overline{\mathscr{E}_n^2(t)} = \frac{1}{2\Pi} \int_0^a \psi_n^2(\tau)\, d\tau . \tag{34}$$

Using Eq. 29 for $\Phi_n^+(S)$, we can replace the summation by an integral when n is very large. For large n the coefficient C_k corresponding to our spectrum can be approximated by

$$C_k(\theta) \approx C^{-2n} \exp\left[\frac{-jn\theta}{2}\right].$$

From Eq. 29

$$\psi_n(\tau) = \begin{cases} \sum_{k=1}^{n} C_k\, e^{p_k \tau} & \tau \geq 0 \\ 0 & \tau < 0 . \end{cases}$$

Thus,

$$\psi_n(\tau) = \begin{cases} \int_{\Pi/2}^{3\Pi/2} \exp\left[\frac{-jn\theta}{2}\right] \exp\left[(a+\tau)\, e^{j\theta}\right] C^{-n}\, d\theta & \tau \geq 0 \\ 0 & \tau < 0 . \end{cases} \tag{35}$$

The functions $\psi_n(\tau)$, as given by Eq. 35, are similar to Fourier coefficients, which have the property of going to zero for large n. This fact is valid for every τ in the interval $[0, a]$; thus we can bound $\psi_n(\tau)$ for n large by

$$|\psi_n(\tau)| < \frac{\Pi\, e^{2\alpha}}{C^{2n}}. \tag{36}$$

Thus by substituting Eq. 36 in Eq. 34 we obtain

$$\overline{\mathscr{E}_n^2(t)} \leq \frac{\Pi}{2}\, \alpha\, \frac{e^{4\alpha}}{C^{4n}}. \tag{37}$$

We can consider n as a measure of the closeness of a nonsingular spectrum to a singular one. In Fig. X-5 the behaviour of $\overline{\mathscr{E}_n^2(t)}$, n, and prediction time into the future is shown.

Note that for $\Phi_n(\omega)$, the error is small until α is of the order of n. Equation 37 shows that for any given prediction time less than or equal to α the error becomes

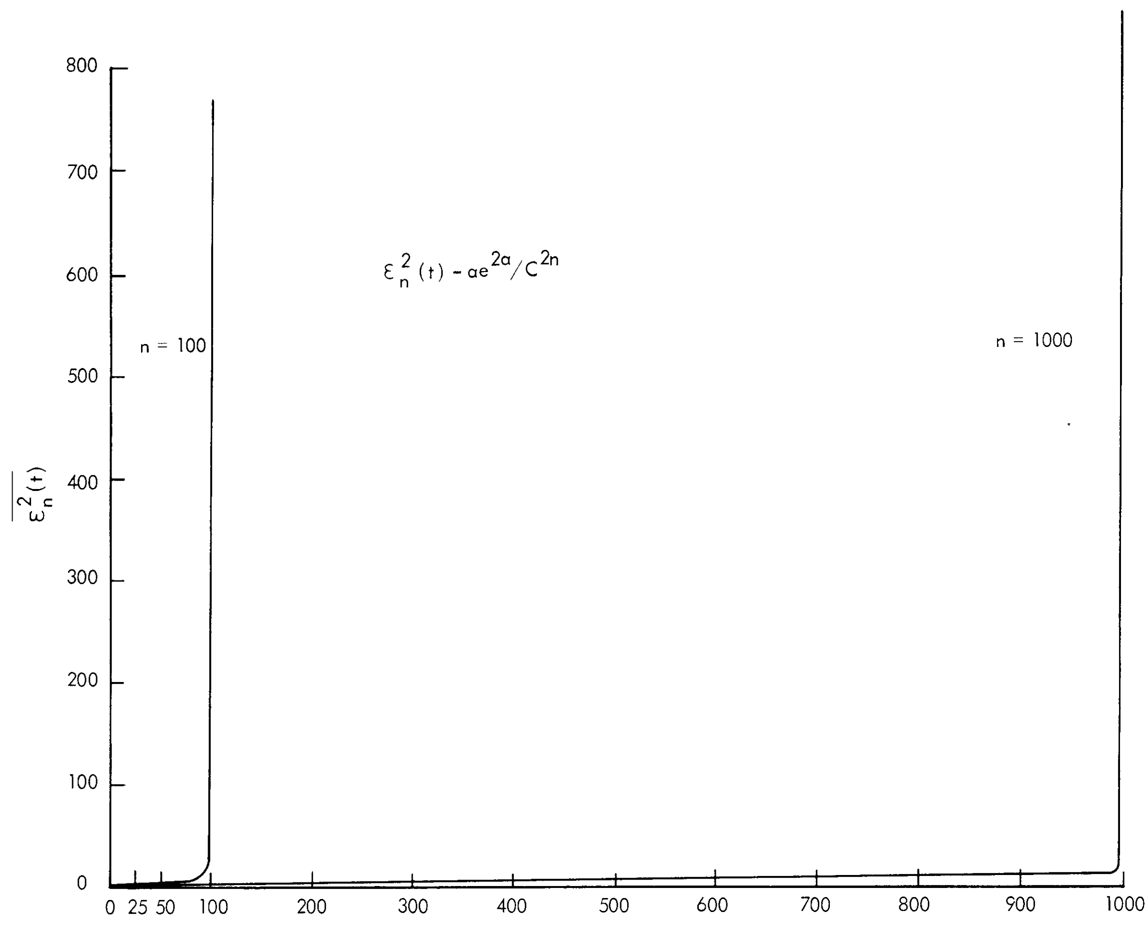

Fig. X-5. Approximate plot of error vs prediction time as a function of the closeness to a singular spectrum (drawn for C = e).

arbitrarily small as n becomes very large. In the limit of infinite n the error becomes zero. This result implies that the optimum predictor is equal to the perfect predictor except over the zero bands of the spectrum which introduce no error whatsoever, as shown by Eq. 10.

The error $\mathscr{E}_n^2(t)$ was small for a less than n because $\psi_n(\tau)$ is small for $\tau < n$ and most of the contribution to the error comes at $\tau \geq n$. Thus for very large n, that is, a spectrum close to being singular, $\psi_n(\tau)$ is small for a long interval $0 \leq \tau \leq n$. This is the interval of good prediction time. This behaviour in $\psi_n(\tau)$ could be seen in other almost singular spectra, as shown in section 4.

This method of predicting a singular spectrum by approximating it arbitrarily closely by a power density spectrum can be applied to other singular spectra of both Classes A and B. For example, the spectra of Fig. X-4b and 4c, $\Phi_{bb}(\omega)$ and $\Phi_{cc}(\omega)$, can be approximated, respectively, by

$$\Phi_{bn}(\omega) = \frac{1}{1+\omega^{2n}} \tag{38}$$

$$\Phi_{cn}(\omega) = 1 + \frac{1}{1+\omega^{2n}} - \frac{1}{1+\left(\frac{\omega}{1+a}\right)^{2n}}. \tag{39}$$

4. The Relation of the Paley-Wiener Integral to the Mean-Square Error of Prediction

It can be shown[5] that the mean-square error of prediction is bounded by

$$\overline{\mathscr{E}_n^2(t)} \leq M\, e^{2a} \exp\left\{\frac{2}{\Pi}\int_{-\infty}^{\infty} \frac{\ln \Phi_{ii}(\omega)}{1+\omega^2}\right\} d\omega \quad . \tag{40}$$

This bound clearly shows that when $\Phi_{ii}(\omega)$ is singular, the Paley-Wiener integral diverges, and the mean-square error of prediction is then zero.

Equation 40 could be used to obtain approximate values of the time of good prediction. Let us consider the spectrum

$$\Phi_n(\omega) = \left(1+\frac{\omega^2}{n}\right)^{-n} \tag{41}$$

which converges to $e^{-\omega^2}$. Applying Eq. 41, we obtain

$$\overline{\mathscr{E}_n^2(t)} \leq M\, e^{2a} \exp\left[-\frac{2n}{\Pi}\int_{-\infty}^{\infty} \frac{\ln(1+\omega^2/n)}{1+\omega^2}\, d\omega\right]$$

$$\sim M\, e^{2a} \exp\left[-\frac{4n}{\Pi}\int_{\omega_o}^{\infty} \frac{\ln(\omega^2/n)}{1+\omega^2}\, d\omega\right]$$

Let $(\omega^2/n) = x$; then

$$\overline{\mathscr{E}_n^2(t)} \leq M\, e^{2a} \exp\left[-\sqrt{n} \int_{\omega_o}^{\infty} \frac{2}{\Pi} \frac{\ln 3x}{x^{3/2}}\, dx\right] = M\, e^{2a}\, e^{-k\sqrt{n}}.$$

For a small compared with $\sqrt{n}$, and for n large, we can see that the error in prediction is very small in agreement with the result given by Wiener[2] which states that the time of good prediction for this wave is of the order of $\sqrt{n}$.

Let us apply the result given by Eq. 40 to another spectrum. Consider the spectrum discussed in section 3, and given by Eq. 27. This spectrum will give a small contribution to Eq. 40 at high frequencies. Thus we need to consider only low frequencies.

$$\overline{\mathscr{E}_n^2(t)} \leq M\, e^{2a} \exp\left[\frac{2}{\Pi} \int_{-\omega_o}^{\omega_o} \frac{\ln \omega^{2n}/(1+\omega^{2n})}{1+\omega^2}\, d\omega\right],$$

where ω_o is some small number. Then

$$\overline{\mathscr{E}_n^2(t)} \leq M\, e^{2a} \exp\left[\frac{2}{\Pi} \int_{-\omega_o}^{\omega_o} \frac{\ln \omega^{2n}}{1+\omega^2}\, d\omega\right] = M\, e^{2a} \exp\left[\frac{4n}{\Pi} \int_{-\omega_o}^{\omega_o} \frac{\ln \omega}{1+\omega^2}\, d\omega\right]$$

$$= M\, e^{2a}\, e^{-kn}. \tag{42}$$

Thus the time of good prediction is of the order of n. This agrees with what we have found in the previous section. Also, note that the bound on the error has a form that is similar to that in the previous section (Eq. 37).

M. Schetzen, A. A. Al-Shalchi

References

1. Y. W. Lee, Statistical Theory of Communication (John Wiley and Sons, Inc., New York, 1960).

2. N. Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series (The Technology Press of Massachusetts Institute of Technology, Cambridge, Massachusetts, and John Wiley and Sons, Inc., New York, 1957).

3. A. A. Al-Shalchi, Prediction of Singular Time Functions, S.M. Thesis, Department of Electrical Engineering, M.I.T., August 1962, pp. 36-38.

4. A. A. Al-Shalchi, Ibid., pp. 29-30.

5. A. A. Al-Shalchi, Ibid., p. 59.

E. ANALOG SIMULATION OF A PHASE-LOCKED OSCILLATOR

A model that describes the behavior of a phase-locked loop is shown in Fig. X-6. The noises $n_1(t)$ and $n_2(t)$ represent the lowpass random processes when the Gaussian input noise is resolved into a component that is in phase with $\theta_1(t)$ and in quadrature

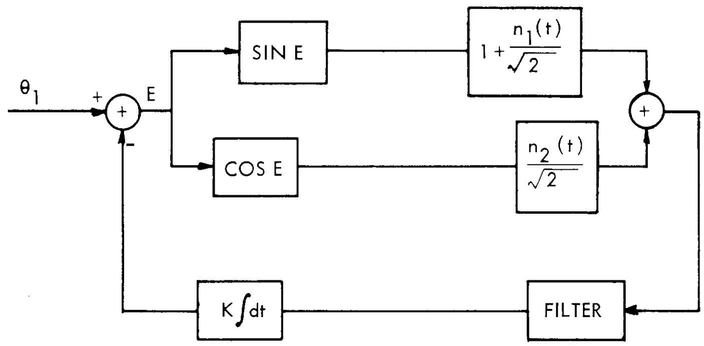

Fig. X-6. Model of a phase-locked oscillator.

with $\theta_1(t)$. The details of the derivation of the model have been given by Van Trees.[1] When $\theta_1(t)$ is constant and the input noise spectrum is symmetrical around the carrier frequency f_c, it is obvious that $n_1(t)$ and $n_2(t)$ are sample functions from independent Gaussian random processes with spectra as shown in Fig. X-7. When the input noise N(t) is white (in practice, when 1/W is small relative to changes in the input phase $\theta_1(t)$), the resultant decomposition still gives independent, Gaussian processes.[2]

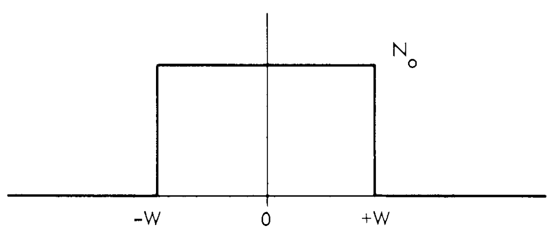

Fig. X-7. Power density spectrum of n_1 and n_2.

The model is valid for all signal-to-noise levels. We are concerned with the threshold behavior. One can show[1] that there exists a noise level that guarantees the nonexistence of any stable equilibrium points.

This level, $N_o > 1/K$, results in an unbounded variance of the phase error, E, as t approaches infinity, and hence this value of N_o is an upper bound on the threshold noise level of the system.

When the model was simulated on an analog computer, the resulting threshold noise levels were found to be less than the predicted upper bound. The details of the simulation and of the experimental results have been given in the author's thesis.[3]

The effects of the two noises acting separately were also studied. In an actual loop the two noises can act separately only if the carrier is amplitude-modulated with

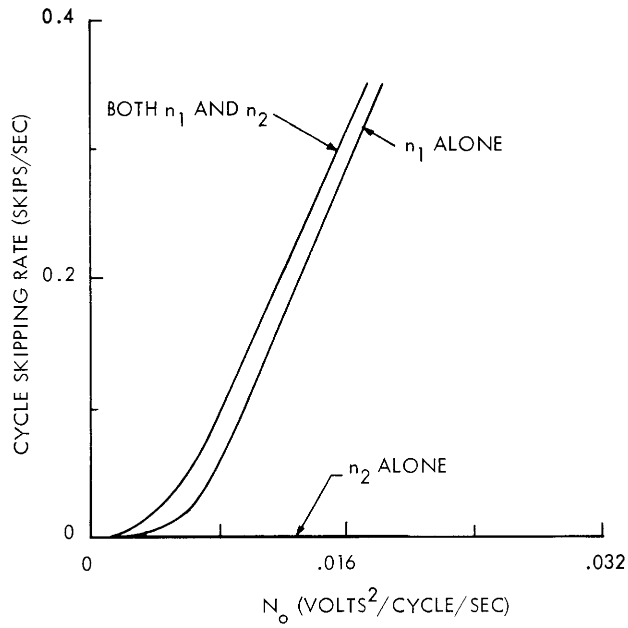

Fig. X-8. Effects of n_1 and n_2 for $\beta/K = 0.3$.

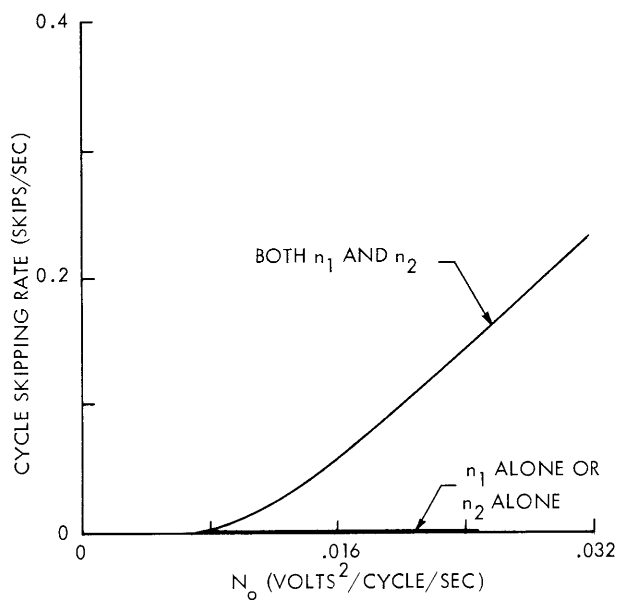

Fig. X-9. Effects of n_1 and n_2 for $\beta/K = 0$.

an independent, Gaussian process. Looking at the noises separately gives some insight into what can be neglected in an analytic model. In Fig. X-8 are shown curves of cycle skipping vs N_o for n_1 acting alone, n_2 acting alone, and n_1 and n_2 acting together, all for $\beta/K = 0.3$. Figure X-9 shows a similar set of curves for $\beta/K = 0$.

We observed that n_2 acting alone did not cause skipping at any value of β/K, and that at higher values of β/K the effect of n_1 acting alone is nearly the same as that for n_1 and n_2 acting together. At lower values of β/K, n_1 did not have as much effect as n_1 and n_2 — in fact, for $\beta/K = 0$, n_1 alone did not cause skipping. This result is obvious, since $n_1(t)$ is multiplying a zero signal.

Similar results could be obtained for a loop with a filter. We expect the corresponding curves to be qualitatively the same as those shown in Fig. X-9.

M. E. Austin

References

1. H. L. Van Trees, Jr., A Threshold Theory for Phase-Locked Loops, Technical Report 246, Lincoln Laboratory, M. I. T., Lexington, Massachusetts, August 22, 1961.

2. H. L. Van Trees, Jr., Self-Synchronization in Bi-Phase Systems (unpublished report).

3. M. E. Austin, Threshold Study of Phase-Locked Loops, S. M. Thesis, Department of Electrical Engineering, M. I. T., June 1962.